

**RAND**

*A Stock-Flow Analysis of  
the Welfare Caseload:  
Insights from California  
Economic Conditions*

*Jacob A. Klerman  
Steven J. Haider*

*DRU-2463-DHHS*

*January 2001*

**DISTRIBUTION STATEMENT A**  
Approved for Public Release  
Distribution Unlimited

**Labor and Population Program**  
**Working Paper Series 01-02**

The RAND unrestricted draft series is intended to transmit preliminary results of RAND research. Unrestricted drafts have not been formally reviewed or edited. The views and conclusions expressed are tentative. A draft should not be cited or quoted without permission of the author, unless the preface grants such permission.

20010510 039

# A STOCK-FLOW ANALYSIS OF THE WELFARE CASELOAD

JACOB KLERMAN AND STEVEN HAIDER  
RAND

1700 Main Street

Santa Monica, CA 90407

klerman@rand.org; sjhaider@rand.org

First Draft: July 1999

Last Revision: January 2001

This work is funded by U.S. Department of Health and Human Services (#99ASPE34A). The opinions and conclusions expressed herein are solely those of the authors and should not be construed as representing the opinions or policy of any of the sponsors or related agencies. Some of the analyses use California administrative data that cannot be released to other researchers without the permission of the State of California. We gratefully acknowledge the expert programming and research assistance of Elizabeth Roth, Laurie McDonald, and Marika Suttorp. We also acknowledge the suggestions of Jay Bhattacharya, V. Joseph Hotz, Darius Lakdawalla, David Loughran, Kathleen McGarry, Terra McKinnish, Regina Ripahn, Robert Schoeni and the seminar participants at the University of Michigan Labor Seminar, the Upjohn Institute for Employment Research, the Institute for Research on Poverty 2000 Summer Workshop, the UCLA Demography Workshop, the University of North Carolina-Chapel Hill Labor Seminar, the RAND/UCLA Labor Seminar, and U.C. Berkeley Welfare Seminar.

## ABSTRACT

During the 1990s, the welfare caseload peaked and then declined by about half. The decline occurred simultaneously with a robust economic expansion and a series of major welfare reforms. This paper reconsiders the methods used in the previous studies to explain these changes. We explicitly model the welfare caseload as the net outcome of past flows onto and off of aid and explore the implications of such a stock-flow perspective for understanding the determinants of the caseload size and its evolution over time. The approach is shown to explain some of the anomalous findings in the literature regarding the effects of economic conditions on the welfare caseload. Then, using administrative data for California, we estimate the effect of the changing unemployment rate on the underlying flows and simulate the impact on the caseload stock. We find that approximately 50 percent of the caseload decline in California can be attributed to the declining unemployment rate. These estimates are substantially larger than the 20 to 35 percent estimates that are obtained from more traditional methods.

During the 1990s, the welfare caseload peaked and then declined by half. The decline occurred simultaneously with a robust economic expansion and a series of major welfare reforms. Applying conventional difference-of-difference models to time-series/cross-sectional data, several studies have estimated the relative importance of the economic expansion and welfare reforms in explaining the caseload decline (e.g., CEA 1997, Levine and Whitmore 1998, CEA 1999, and Ziliak et al 2000). These studies have reached widely varying conclusions regarding the cause of the caseload decline. The different conclusions appear to be due to different specifications for the relationship between the current welfare caseload and lags of the explanatory and dependent variables (Figlio and Ziliak 1999, and CEA 1999). However, none of the studies explicitly consider the source of these relationships.

This literature has developed nearly independently of a large literature that examines the flows onto and off of welfare (e.g., Hutchens 1981, Bane and Ellwood 1986, and Hoynes 2000). In this paper, we directly consider the implications of viewing the welfare caseload as the net outcome of past flows onto and off of welfare. This stock-flow approach suggests a source for the strong dependence of the caseload on lags of the explanatory variables found in the previous literature. Furthermore, it suggests that the conventional models are mis-specified and this mis-specification can explain the disparate results from the previous literature.

Beyond suggesting a critique of the existing literature, this stock-flow perspective offers an alternative estimation strategy. Specifically, we can directly estimate the underlying flow relationships and then simulate the implied impact on the caseload stock (e.g., Heckman and Walker 1992 and Moffitt and Rendall 1995). Such an approach is not feasible with the available national data.<sup>1</sup> Instead, we use California administrative data to estimate the dependence of the flows on economic conditions. Because these data are only for one state, we are not able to distinguish the impact of policy changes from

---

<sup>1</sup> Most previous caseload studies use national repeated cross-sectional data. Moffitt (1993) shows that underlying flow parameters cannot be identified from repeated cross-sectional data. The longitudinal survey data that are available do not have sufficiently large sample sizes to estimate the type of models we develop. In addition, survey data appear to suffer from serious data quality problems, such as under-reporting of program participation and seam bias (e.g., Bavier, 2000).

more general time effects. However, there is sufficient variation to precisely estimate the effect of changing economic conditions, and our results suggest that approximately half of the caseload decline in California can be attributed to changing economic conditions, as measured by the unemployment rate.<sup>2</sup> These estimates are substantially larger than the 20 to 35 percent estimates that are obtained from more traditional methods.

The results from this paper have implications beyond understanding the impact of the unemployment rate on California's welfare caseload. First, the empirical results for California stock regressions are very similar to results presented in the national literature; thus, it is reasonable to expect similar results would be found if national data were available. Second, the strong empirical support for the stock-flow model suggests that a simple differencing strategy applied to the welfare stock data will give mis-leading results when evaluating any underlying change, including policy change. Finally, stock-flow concerns are likely to be important when evaluating many other economic outcomes. For example, Schoeni (2000) demonstrates that the studies examining the changing foodstamp caseload exhibit similar patterns (e.g., Wallace and Blank 1999; Currie and Grogger 2000; and Figlio, Gunderson, and Ziliak, and 2000).

The balance of the paper is organized as follows. Section 1 develops the basic stock-flow framework. We briefly discuss our data in Section 2. Section 3 presents conventional stock regressions based on the California data. Section 4 presents new estimates on the role of the economy that are based directly on the stock-flow framework. We discuss our conclusions in Section 5. Appendix A provides additional analytical results and Appendix B provides further details on our data set construction.

## 1. A STOCK-FLOW MODEL OF THE CASELOAD

In this section, we first outline a conceptual model of welfare that motivates modeling entry and continuation separately rather than simply modeling the aggregate caseload.

---

<sup>2</sup> For a more comprehensive assessment of the caseload decline in California, see the series of reports from RAND's statewide evaluation (<http://www.rand.org/CalWORKs>) and MaCurdy, Mancuso, and O'Brien-Strain (2000).

This conceptual model is the basis of our stock-flow approach. We then specify an extremely simple model in which the welfare caseload is the result of underlying flows that follows from straightforward accounting identities. We use the model to consider the implications of our stock-flow approach for the conventional aggregate regressions. We then propose a new estimation strategy that is based on re-writing the stock-flow model as a standard Markov process.

*A Conceptual Model of Welfare Receipt.* Consider a conceptual model of welfare use in which individuals choose to be on welfare if the utility of receiving welfare is greater than the utility of working (Moffitt, 1983). Because local labor market conditions will affect the wage offer an individual receives, local labor market conditions will directly affect whether individuals choose to receive aid. For example, as the local economy deteriorates, an individual's best job offer (i.e., wage offer) will weaken and the individual will be more likely to choose to receive welfare.

A simple probability model of welfare receipt would follow directly from this conceptual model if the process remains static, i.e., if welfare receipt in the prior period does not affect welfare receipt in the current period. Under such circumstances, we could posit that the probability an individual receives aid is a function of explanatory variables such as current economic conditions, policy variables, and general time effects. Estimation could proceed using conventional binary choice models, with either individual or grouped data. This estimation strategy is similar that used in many previous studies in which the log of the per capita caseload is modeled directly.<sup>3</sup>

Consider the case, however, when welfare receipt is not static but dynamic. In other words, suppose that the probability that an individual receives welfare this period depends on whether and/or how long the person received welfare last period. Many different underlying processes would cause such a dynamic model to arise. For example,

---

<sup>3</sup> Studies in the recent strand of literature include CEA (1997), Levine and Whitmore (1998), CEA (1999), Figlio and Ziliak (1999), and Ziliak et al (2000). Other studies that use a similar methodology include Blank (1997), Blank and Wallace (1999), Moffitt (1999), and Schoeni and Blank (2000). Congressional Budget Office (1993) provides a list of earlier studies that examine the effect of economic conditions on the caseload, with a focus on forecasting the caseload. Moffitt (1992) provides a more general review of the literature on welfare.

fixed costs of entering or exiting welfare would cause welfare receipt to depend on previous welfare receipt. Similarly, if job offers arrive stochastically (e.g., Lippman and McCall 1976), then an improved labor market will lead to an increase in the exit rate from welfare, but the full effect on the stock will not be felt for several periods. In addition, human capital depreciation or negative stigma associated with welfare could induce dependence with respect to the amount of time on and off of aid; this type of state dependence is often referred to as “duration dependence.” Moreover, there is a large empirical literature that explicitly incorporates such dependence in welfare receipt by modeling the entry and exit to the welfare caseload separately.<sup>4</sup> These studies show clear empirical evidence that welfare receipt is dependent on previous welfare receipt, even conditional on covariates.

When the welfare receipt follows a dynamic process, the flows (i.e., entry onto and continuation on welfare) by definition follow different processes. For example, we could posit that the probability an individual enters welfare (i.e., receives welfare conditional on not receiving welfare in the previous period) is a function of explanatory variables such as current economic conditions, policy variables, and general time effects. A similar but separate model could be posited for continuing on welfare. The impact of a change in an explanatory variable on the caseload stock would then operate through changes in both of the flows, with the particular relationship given by accounting identities. We refer to such a model as a stock-flow model.

*Uncovering the Impact of Economic Conditions on the Caseload Stock.* To explore the implications of a stock-flow approach for modeling the stock directly, we begin with an extremely simple case. Suppose that welfare spells last either one or two periods. In this case, today’s caseload is equal to the sum of today’s entrants and the fraction of yesterday’s entrants that are still on aid. Formally, letting  $N_t$  be the caseload in period  $t$ ,  $E_t$  be the number of new entrants onto welfare in period  $t$ , and  $c_t$  be the continuation rate

---

<sup>4</sup> Hutchens (1981) examines separately entry onto and exit off of welfare. Bane and Ellwood (1986) and Blank (1989) examine the determinants of welfare spell lengths. Moffitt (1992) provides a useful review of the earlier studies. Hoynes and MaCurdy (1994) examine how changes in program generosity affect spell length. Gittleman (1999) and Hoynes (2000) are more recent studies of welfare dynamics.

(the probability of receiving aid in period  $t$  if one received aid in period  $t-1$ , i.e., one minus the conventional hazard rate), then the caseload in period  $t$  can be expressed as,

$$(1) \quad N_t = E_t + c_t E_{t-1}.$$

Furthermore, suppose that there is a homogenous population of individuals at risk of going on welfare,  $P_t$ , and that the rate of entry per person at risk is  $e_t$ . Then, the number of entrants can be written as,

$$(2) \quad E_t = e_t P_t.$$

Now, assume that the entry rate,  $e_t$ , and the continuation rate,  $c_t$ , are functions only of current economic conditions  $Y_t$ , i.e.,  $e_t = e(Y_t)$  and  $c_t = c(Y_t)$ . Substituting these quantities into equation 1 and dividing through by the population, we have

$$(3) \quad n_t = e(Y_t) + c(Y_t)e(Y_{t-1}),$$

where  $n_t$  is the caseload per capita ( $n_t = N_t/P_t$ ).<sup>5</sup> Finally, suppose we further assume that the entry rate and the continuation rate are linear functions of contemporaneous economic conditions,

$$(4) \quad e_t = \alpha_e + \beta_e Y_t + \varepsilon_{et},$$

$$(5) \quad c_t = \alpha_c + \beta_c Y_t + \varepsilon_{ct}.$$

Substituting equations 4 and 5 into equation 1, we obtain the following expression,

---

<sup>5</sup> This derivation implicitly assumes that the population at risk of going on aid (the total population less those already on aid) and the total population are identical. This notation is approximately correct whenever the fraction of the population on aid is relatively small. We choose this notation for expositional convenience. A Markov representation would precisely distinguish between the total population and the population at risk. However, the resulting expressions would be much more complicated, obscuring the implications of our stock-flow approach for the conventional stock regressions. Below, we use the Markov representation for our micro-level empirical results.



$$\begin{aligned}
n_t &= e_t + c_t e_{t-1} \\
&= (\alpha_e + \beta_e Y_t + \varepsilon_{et}) + (\alpha_c + \beta_c Y_t + \varepsilon_{ct}) (\alpha_e + \beta_e Y_{t-1} + \varepsilon_{et-1}) \\
(6) \quad &= (\alpha_e + \alpha_c \alpha_e) + (\beta_e + \alpha_e \beta_c) Y_t + (\alpha_c \beta_e) Y_{t-1} + (\beta_e \beta_c) Y_{t-1} Y_t \\
&\quad + (\alpha_c + \beta_c Y_t) \varepsilon_{et-1} + (\alpha_e + \beta_e Y_{t-1}) \varepsilon_{ct} + \varepsilon_{ct} \varepsilon_{et-1} \\
&= \pi_0 + \pi_1 Y_t + \pi_2 Y_{t-1} + \pi_3 Y_t Y_{t-1} + v_t.
\end{aligned}$$

This simple model suggests several observations that are more general. First, even when the flows (the entry rate and the continuation rate) are functions only of current economic conditions, the caseload  $n_t$  is a function of current and lagged economic conditions (see equation 3). Second, rather than entering as a conventional distributed lag, economic conditions enter in a non-linear, interactive form (equation 6). Third, these relationships are in levels of the per capita caseload, not logs, as specified in the time-series/cross-section literature.<sup>6</sup> Finally, with the additional assumptions that the error terms in equations 4 and 5 are independent of economic conditions and serially uncorrelated, linear regression techniques could be used to estimate a reduced-form model of the welfare stock and economic conditions (equation 6), and non-linear regression techniques could be used to estimate the structural parameters (i.e., the parameters of the flow relationships).

The expressions become more complicated as the model is extended to allow for welfare spells that last for more than two periods, but analogous observations continue to apply. We organize these observations into three simple propositions, and then discuss the implications of the propositions to the conventional stock models.

To motivate the first proposition, we first extend the basic caseload identity (equation 3) to allow for individuals to be on aid for  $K$  periods but maintain the assumption that the continuation rate is invariant with duration,

---

<sup>6</sup> Our general point is that the underlying accounting identities provide significant structure to the problem. The exact structure for the stock regression is dependent on the functional forms that are assumed for the underlying flow regressions. Some aspects of this structure, such as the precise distinction between levels and logs, are not robust to alternative specifications. However, other aspects such as the need for many lags and the interactive relationships are robust. We exploit all of the available structure here to be internally consistent.

$$(7) \quad n_t = e(Y_t) + \sum_{k=t-K}^{t-1} e(Y_k) \prod_{j=k+1}^t c(Y_j).$$

The first proposition is simply an observation about the structure of this expression:

**Proposition 1.** If the entry rate and the continuation rate are functions only of contemporaneous economic conditions, then the per capita caseload will be a non-linear function of lagged economic conditions equal to the longest period individuals are on aid.

This proposition has three implications for the conventional approach of assessing why the caseload has declined (e.g., CEA 1997; Levine and Whitmore 1998; and CEA 1999). The first implication implies that the regressions require as many lags of economic conditions as the number of periods people remain on aid. The intuition behind this implication is that the caseload may be large today because economic conditions were particularly bad a few years ago, and some of these individuals have remained on aid. The empirical importance of such concerns depends on how long individuals remain on aid after they enter. Bane and Ellwood (1994) calculate that over 25 percent of welfare spells last over 5 years, and over 65 percent of the persons on welfare at a particular point in time have been on welfare for at least 5 years.

Second, the proposition also implies that interactions between the economic condition lags are likely to be important in models that use stock data. Again, the intuition is straightforward: bad economic conditions a few years ago will have a much different impact on today's caseload depending on whether or not the intervening years had good or bad economic conditions.

The third implication of the first proposition is that the reduced-form stock relationship and underlying structural parameters could be estimated if additional assumptions were made about the functional form of the entry and continuation rates. For example, we demonstrated this was possible assuming the flows were linear in economic conditions and  $K = 2$  (see equation 6). Recall that such a model should be specified in levels, not logs. Precise estimation of the parameters would require a time-series sufficiently longer than  $K$  periods. In practice, the available time series is short relative to length of time individuals remain on aid.

The second proposition states that a second and simpler method of estimating equation 7 is available, if we maintain the assumption that there is no duration dependence. In particular, we show in the appendix that the inclusion of a lagged *dependent* variable, as well as an interaction with the contemporaneous economic conditions, can approximately identify the underlying structure of the model. We state this as a second proposition.

**Proposition 2.** If the entry rate and the continuation rate are linear functions only of contemporaneous economic conditions, then a lagged dependent variable can approximately recover the reduced-form or structural relationship between the caseload stock and economic conditions.

The intuition for this proposition is straightforward: because the continuation rate is assumed to be invariant to the time on aid, the lagged aggregate caseload contains all necessary information about past economic conditions. Thus, we can recover the underlying structure by estimating the model,

$$(8) \quad n_t = \theta_1 + \theta_2 Y_t + \theta_3 n_{t-1} + \theta_4 (n_{t-1} Y_t) + v_t.$$

We derive this expression in the appendix. Although such a justification is not given in Ziliak, et al (2000) or Figlio and Ziliak (1999), it could be used to provide a structural interpretation for their inclusion of a lagged dependent variable. However, such a structural model should be estimated with the dependent variable specified in levels, not logs, and include the indicated interaction.

The simple solution of including a lagged dependent variable is not general. It requires that there be no duration dependence (i.e., that the continuation rate does not depend on the length of time on aid), an assumption that is clearly rejected by the data (e.g., Bane and Ellwood, 1994). To demonstrate this claim, we continue to assume that the continuation rate is still a function of current economic conditions, but we let the continuation rate vary with the duration of aid receipt. Define  $c_t^k$  to be the continuation rate in period  $t$  for welfare spells that began  $k$  periods ago ( $c_t^k = c^k(Y_t)$ ). If individuals could be on aid for  $K$  periods, then the aggregate caseload would be,

$$(9) \quad n_t = e(Y_t) + \sum_{k=t-K}^{t-1} e(Y_k) \prod_{j=k+1}^t c^{k-j}(Y_j).$$

As before, this expression implies that to identify the underlying structure of the model, it is necessary to include  $K$  lags of economic conditions and a complete set of interactions (up to the  $K$ th-order). Furthermore, if the continuation rate varies with duration, including a lagged *dependent* variable is no longer sufficient to obtain unbiased estimates. If only a lagged dependent variable were included, then that coefficient would measure the *average* effect of economic conditions across the caseload. This average coefficient will depend on the distribution of spell lengths among those currently on welfare, which in turn depends on the history of economic conditions. We state this observation as a third proposition.

**Proposition 3.** If the effect of economic conditions on the continuation probability depends on the length of time on aid, then: (A) The coefficient on the lagged dependent variable will be a linear combination of the duration-dependent continuation probabilities. And (B), the weights of the linear combination will vary with economic conditions.

This proposition follows from writing the continuation rate as a weighted average of the duration specific continuation rates,

$$(10) \quad n_t = e_t + [c_t^1 \lambda_{t-1}^1 + c_t^2 \lambda_{t-1}^2 + \dots + c_t^K \lambda_{t-1}^K] n_{t-1}.$$

where  $\lambda_{t-1}^k$  is the share of the caseload in  $t-1$  that began  $k$  periods previous; see the appendix for details. The implication of this proposition is that the coefficient on the lagged dependent variable will vary with the history of economic conditions. Thus, with duration dependence, simply including a lagged dependent variable and an interaction with the contemporaneous unemployment rate will not recover the underlying structure.

In summary, the relationship between the caseload stock and economic conditions becomes very complicated under plausible assumptions, even when the relationship between the flows and economic conditions are extremely simple. Moreover, a stock-flow model has direct implications for how stock regressions should be specified. These

observations follow simply from writing down the accounting identities that describe how the caseload evolves over time.

*Estimating the Change in the Stock from the Changes in Flows.* Not only does the stock-flow model suggest an explicit functional form for when stocks are analyzed directly, but it also offers an alternative estimation strategy for when flow data are available. Specifically, we can estimate the flow relationships directly and then simulate the responsiveness of the aggregate caseload to changes in economic conditions. Such methods have been used in several other literatures (e.g., Heckman and Walker, 1992; Moffitt and Rendall, 1995).

To develop such an empirical model, we re-write the stock-flow model as a conventional Markov equation.<sup>7</sup> Allowing for duration dependence in the process for leaving welfare, the Markov equation can be represented as,

$$(11) \begin{bmatrix} S_{r,1,t} \\ S_{r,2,t} \\ S_{r,3,t} \\ \vdots \\ S_{r,\bar{k}-1,t} \\ S_{r,\bar{k},t} \\ S_{n,t} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 & e(Y_t) \\ c^1(Y_t) & 0 & \dots & 0 & 0 & 0 \\ 0 & c^2(Y_t) & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \dots & c^{\bar{k}-1}(Y_t) & c^{\bar{k}}(Y_t) & 0 \\ 1-c^1(Y_t) & 1-c^2(Y_t) & \dots & 1-c^{\bar{k}-1}(Y_t) & 1-c^{\bar{k}}(Y_t) & 1-e(Y_t) \end{bmatrix} \begin{bmatrix} S_{r,1,t-1} \\ S_{r,2,t-1} \\ S_{r,3,t-1} \\ \vdots \\ S_{r,\bar{k}-1,t-1} \\ S_{r,\bar{k},t-1} \\ S_{n,t-1} \end{bmatrix}$$

where  $S_{rkt}$  is the number of individuals who are receiving aid for the  $k$ th consecutive period at time  $t$  and where  $S_{n,t}$  is the number of individuals not on aid.<sup>8</sup> To address initial conditions problems discussed in more detail below, this formulation assumes that the continuation rate varies through period  $\bar{k}$  and is constant thereafter.

We can represent the Markov equation more compactly as,

<sup>7</sup> This Markov formulation relaxes the simplification that was discussed in footnote 5. In particular, the population at risk of entering aid is explicitly defined to be the total population less those currently on aid.

<sup>8</sup> Haider, Klerman, and Roth (2001) estimate a more general specification in which entry and re-entry onto aid are allowed to differ and re-entry is allowed to vary by the amount of time off aid. Estimates of that more complicated model yield results quantitatively similar to the simpler model discussed and estimated in this paper.

$$(12) \quad \underset{(Q \times 1)}{S_t} = \underset{(Q \times Q)}{M(Y_t, \theta)} \underset{(Q \times 1)}{S_{t-1}},$$

where  $S_t$  is a vector that contains the number of individuals in each of  $Q$  "states" and  $M$  is the transition matrix between the states. The transition matrix  $M$  depends on economic conditions  $Y_t$ , a parameter vector  $\theta$ , and possibly other explanatory variables.

We can use this equation to simulate the impact of economic conditions on the caseload stock implied by the underlying flow relationships. We first estimate models for the flows (i.e., the entry rate and the continuation rate) to obtain estimates of the parameter vector  $\theta$ . Then, for any arbitrary specification for economic conditions  $\tilde{Y}_t$ , we can calculate the implied transition matrix  $M$  and simulate the caseload for the following period. Thus, given an initial stock  $S_t$  and any arbitrary path for economic conditions,  $\{\tilde{Y}_{t+j}\}$ , we can simulate the future stock in period  $t+k$  as,

$$(12) \quad S_{t+k} = \left( \prod_{j=1}^k M[\tilde{Y}_{t+j}, \theta] \right) S_t$$

This expression can be used to explore the effect of the economy by specifying an explicit counterfactual history. For example, below we simulate the implied path of the caseload for a history in which the unemployment rate had remained at its peak. Similar methods potentially could be used to examine the impact of any explanatory variable, including the impact of policy.

## 2. THE DATA

Directly estimating the stock-flow model requires panel data on individuals, across sufficiently varying explanatory variables (e.g., economic or policy conditions), for a sufficiently long time period, and for a sufficiently large sample. Appropriate national data do not appear to be available. There exists no individual-level national administrative data, and the available panel surveys appear to record welfare receipt with considerable error and the sample sizes appear to be too small to estimate the transitions

that are the focus of this study (e.g., the Panel Study of Income Dynamics or the Survey of Income and Program Participation).

Instead, we estimate our stock-flow models using administrative data from California's Medi-Cal Eligibility Determination System (MEDS).<sup>9</sup> The MEDS provides a monthly roster of welfare participants in California from 1987 to the present.<sup>10</sup> In addition, the MEDS is recorded as part of an on-going administrative process so biases associated with self-reports are not present. Moreover, there is significant diversity in economic conditions across California's counties. This diversity allows us to use an identification strategy that is similar to that used in the national literature (i.e. variation across the 50 states) to examine the role of economic conditions across California's 58 counties. Finally, California is a large state with over twenty percent of the U.S. welfare caseload and over ten percent of the U.S. population; caseload trends in California comprise a significant share of the caseload trends in the U.S.

The MEDS data have one major disadvantage in that they cover only a single state. Although the impact of state-level policy variation is one of the motivations in the national literature, state-level policy only varies temporally in our data and we do not have a plausible identification strategy to separate policy changes from more general secular changes.<sup>11</sup> Therefore, we concentrate on the impact of changing economic conditions.

Appendix B discusses the details of our analysis file. In brief, we construct an analysis file by drawing a stratified random sample of approximately 3 percent of the full MEDS. Consistent with our focus on the effect of county level variation in the economy, the stratified sample is chosen to yield approximately equal number of observations in

---

<sup>9</sup> Hoynes (2000) uses a different extract from the same underlying database.

<sup>10</sup> Throughout this paper, we use the term "welfare" to refer to the Aid For Families with Dependent Children (AFDC) program that was changed to the Temporary Assistance for Needy Families (TANF) program. The programs provide financial assistance to needy families (usually headed by a single mother) with children. Such participation can be identified with the MEDS because welfare recipients are categorically eligible for Medi-Cal (the California implementation of Medicaid), and such eligibility is designated in the database.

each of California's counties. This scheme results in an analysis file that contains a sample of 282,381 people who received cash assistance, comprising 487,641 spells and 10,966,420 person-months, during the years 1989 to 1998 (our eventual sample period).

We present the aggregate monthly caseload estimated from our MEDS-based analysis file, the official state caseload counts (based on county-level CA237 reports), and the unemployment rate in Figure 1. The figure shows that the MEDS tracks the official caseload count well.

Figure 1 also demonstrates that the paths of the caseload and the unemployment rate suggest a role for the economy in explaining the caseload decline. In particular, the caseload increased during the early 1990s and then declined during the latter 1990s, similar to the trend for the United States as a whole. At the peak of the welfare caseload in March 1995, there were approximately 2.7 million people receiving AFDC/TANF in California. In the last month of our sample period, December 1998, there are only 1.9 million people on aid, representing a decline from the peak of 31 percent. Because the population increased during the same time frame, the per capita caseload declined by over 33 percent. Turning to the unemployment rate, the figure shows that the unemployment rate increased then decreased, following a similar pattern to that of the caseload. In particular, the unemployment rate declined from 10 percent at its peak to 6 percent at the end of our sample period. We conclude by noting that the unemployment rate appears to lead the caseload as is implied by our stock-flow model; in particular, changes in the unemployment rate today affect the flows immediately, but the changes in the stock build up over time.

The basis of the stock flow model is that the entry rate and continuation rates differ (i.e., that there is dependence with respect to previous welfare receipt) and that these rates vary with economic conditions. Table 1 presents the average monthly entry rate and continuation rate for two-year intervals between 1989 and 1997. These tabulations reveal

---

<sup>11</sup> The national stock literature uses time fixed effects in their specifications, and we adopt a similar specification for our flow models; again, this allows us to isolate the impact of moving to a stock-flow model as compared to the conventional models in the literature.



several important characteristics of the data. First, the levels of the entry and continuation rates in Table 1 are quite different, supporting the claim that dynamics are important. The average monthly entry rate for those who were not on welfare in 1989-90 was 0.0032 and average monthly continuation rate for those who were on welfare for 2 to 5 months was 0.938.<sup>12</sup> Second, both the entry and continuation rates are counter-cyclical. The entry rate increased during the recession of the early 1990s (from 0.0032 to 0.0037) and then declines during the recovery (back to 0.0024). Similarly, for all durations, the continuation rate increased then decreased. Both the entry and continuation rate patterns would cause the aggregate caseload size to vary counter-cyclically with economic conditions. Finally, the continuation rate is higher for individuals who have been on aid longer (i.e., there is negative duration dependence).

### 3. RESULTS FROM CONVENTIONAL STOCK REGRESSIONS

In Section 1, we derived several implications of a stock-flow model for the conventional stock regressions. In this section, we discuss the literature that has estimated conventional stock regressions using national data and present conventional stock regressions for California's counties. We show that the California results are quantitatively and qualitatively very similar to those estimated with national data. Moreover, we demonstrate that many of the implications of the stock-flow model are borne out by the data.

*Results from the Conventional Stock Regressions.* The national studies that have estimated conventional stock regressions have come to widely varying conclusions. The recent literature begins with a Council of Economic Advisers analysis of the effect of pre-PRWORA welfare waivers on the aggregate caseload with administrative data (CEA 1997). That study attributes 44 percent of the decline in the caseload from 1993-96 to economic conditions and 31 percent of the decline to welfare waivers. Ziliak, et al

---

<sup>12</sup> For this and later analysis, we restrict the population at risk to be all individuals under the age of 50. We include men in the analysis because our analysis includes the smaller AFDC-Unemployed Parent program (AFDC-UP), which provides welfare benefits to two parent families in which the husband has recently lost a job.

(2000), examining exactly the same question, attributes nearly two-thirds of the same decline to economic conditions and nothing to welfare waivers. Figlio and Ziliak (1999) attempt to reconcile the differences between CEA (1997) and Ziliak, et al (2000) and conclude that the primary reason for different results rests with differences in modeling the “dynamics”, where dynamics refers to whether lagged dependent variables are included and whether the models are estimated in levels or differences. CEA (1999), updating the previous CEA study, finds that including a second lag of the unemployment rate reduces the role of the economy from 42 percent to 26 percent.

Mimicking the national literature, we estimate the following model with annual California data,

$$(13) \quad \ln n_{jt} = \alpha + Y_{jt}\beta + \gamma_t + \delta_j + \varepsilon_{jt}$$

where, for county  $j$  and year  $t$ ,  $n_{jt}$  is the welfare recipients per capita,  $Y_{jt}$  is a vector of the unemployment rate and its lags,  $\gamma_t$  is a time fixed effect, and  $\delta_j$  is a county fixed effect.<sup>13</sup> We will refer to equation 13 without a lagged dependent variable (similar to the CEA specification) as a “static stock” model and equation 13 with a lagged dependent variable (similar to the Ziliak specification) as a “dynamic stock” model. We estimate all regressions on data for the years 1988-98 with ordinary least squares (OLS), using data aggregated to the county-year.

Table 2 presents results for the static stock model (i.e., without a lagged dependent variable) in columns 1 to 4 and results for the dynamic stock model (i.e., with a lagged dependent variable) in columns 5 to 8. The first model in each group only includes the contemporaneous unemployment rate. Successive models include additional lags of the unemployment rate. The results for the both sets of regressions are similar qualitatively and quantitatively to those calculated using national data.

---

<sup>13</sup> Some of the aggregate literature uses the log of the per capita caseload as the dependent variable rather than the log of the per capita recipients. Based on our comparisons (presented in earlier drafts of the paper and available from the authors) and those of others (Figlio and Ziliak, 1999), the results are almost identical when using either dependent variable. We analyze persons rather than cases to retain a well-defined, longitudinal unit of analysis.

First, consider the static stock results presented in columns 1 to 4 of Table 2. The coefficient in model 1 implies that a 1 percent increase in the unemployment rate is associated with a 2.2 percent increase in the welfare caseload. To compare the results across the various models, we sum the coefficients to calculate a “long-run elasticity” of the caseload with respect to a permanent change in the unemployment rate. The long-run elasticity increases as additional lags are added, from 2.5 percent (no lags), to 3.5 percent (one annual lag), to 4.7 percent (two annual lags), and then to 5.9 percent (three annual lags). These estimates on California data with a shorter time period match the national literature quite well; for example, a comparable calculation from the national literature for two annual lags is 5.4 percent (see CEA 1999, model 1 in Table 2).

Previous studies have described the economic significance of these coefficients by calculating the percent of the caseload decline that can be explained by economic conditions. We provide similar calculations for California for the period 1994 (the peak in the annual California caseload) to 1998. We calculate the percentage of the decline that can be explained by the economy in the static stock model,  $\Delta_{econ}^s$ , as

$$(14) \quad \Delta_{econ}^s = \frac{\exp[(Y_{1998} - Y_{1994})\hat{\beta}] - 1}{(n_{1998} - n_{1994})/n_{1994}},$$

where  $Y_t$  is the unemployment rate for California and  $n_t$  is the total caseload in California, both in year  $t$ . This expression, as well extensions to incorporate other explanatory variables, is derived in Appendix A. For no lags, the model implies 20 percent of the caseload decline can be attributed to the declining unemployment rate. This percentage increases to 36 percent when one or two lags are included but then declines back to 20 percent when three lags are included. This decline in the amount explained by the economy occurs despite the fact that the long-run elasticity is monotonically increasing with the addition of lags. Both this sensitivity and the divergence in the patterns with the inclusions of more lags (between the long run elasticity and the percentage explained) are also found in the national literature.

This sensitivity in results is likely related to a peculiar empirical regularity. For the inclusion of different number of lags of the unemployment rate, the coefficient on the longest included lag is quantitatively the largest and highly statistically significant in each, while the shorter lags tend to be close to zero or even wrong signed.<sup>14</sup> This is true for the models in columns 1 to 4 presented here, as well as for CEA (1999) models that include one and two annual lags and some of the Figlio and Ziliak (2000) dynamic models that include four annual lags.

The relationship between the longest lag being empirically the largest and the sensitivity of results appears to arise because increasing the number of included lags changes the primary unemployment rate change included in the calculation for the role of the economy. For example, in the calculations presented in Table 2, we compute the role for the economy for the period 1994 to 1998. Given that the longest included lag has the largest coefficient in all of the models, a model with two lags is primarily a model of the unemployment rate change from 1992 to 1996, while a model with three lags is primarily a model of the unemployment rate change from 1991 to 1995. The net unemployment rate change for 1992 to 1996 is -2.0 percentage points (9.4 to 7.4 percent), but the net unemployment rate change for 1991 to 1995 is +0.1 percentage points (7.7 to 7.8 percent).<sup>15</sup> Thus, the unemployment rate change used in the latter model is much smaller, and the role for the economy is calculated to be much smaller. Therefore, beyond being an empirical anomaly, the largest coefficient on the longest lag appears to induce a disturbing sensitivity to the empirical calculations.

---

<sup>14</sup> Wrong-signed based on the a priori belief that a high unemployment rate should lead to a high welfare caseload. Previous authors have suggested that lagged unemployment might be more important than contemporaneous unemployment "because welfare recipients are likely to be the last ones hired during an economic recovery and thus may not instantaneously move from welfare to work" (see Figlio and Ziliak, 1999, p. 32-33). However a natural extension of this argument is that welfare recipients are also the *first* to be fired during economic downturns; over the business cycle, these effects would tend to be offsetting. Moreover, the fact that the longest lag is always most important, even when using three annual lags (results presented below), casts doubt on such an argument.

<sup>15</sup> Exactly the same argument can explain the sensitivity of the CEA results. CEA (1999) reports that the amount of the 1993 to 1996 caseload decline due to economic conditions is 44 percent with one lag but declines to 22 percent with two lags. The change in the national unemployment rate from 1992 to 1995 is 2.3 percentage points (7.9 to 5.6 percent). The change in the national unemployment rate from 1991 to 1994 is 0.8 percentage points (6.9 to 6.1 percent).

In columns 5 through 8 of Table 2, we present results from a dynamic stock model, similar to Ziliak, et al. (2000) and Figlio and Ziliak (1999). Although the net impact of the unemployment coefficients is remarkably similar with or without the lagged dependent variable, the coefficient on the lagged dependent variable is close to one and highly significant. Because of the large lagged dependent variable coefficient, the long-run impact of the unemployment rate is substantially higher.<sup>16</sup>

Just as in the static stock case, we can describe the economic significance of these coefficients by calculating the percent of the caseload decline that can be explained by the unemployment rate. For the dynamic case, this quantity takes into account the feedback nature of including a lagged dependent variable (see the appendix for the explicit development). Using California data, the declining unemployment rate explains 62 percent of the caseload decline between 1994 and 1998 for the dynamic model that includes no lags and increases to 73 percent when one unemployment rate lag is included, and declines to 41 percent when three unemployment rate lags are included. These results are quantitatively similar to the calculations with national data (see Figlio and Ziliak, 1999). Although the last lag is not always the most important in these specifications, the percentage of the decline attributable to the declining unemployment rate is still highly sensitive to the number of lags chosen.

*The Implications of the Stock-Flow Model for the Stock Regressions.* Referring back to columns 1 to 4 of Table 2, the static stock model estimates varied between 20 and 35 percent, depending on the number of lags that were used in the model. Our analytic results for the stock-flow model suggest that these models suffer from an omitted variables bias both because they truncate necessary lags and that they ignore important interactions that are implied by the underlying process. This omitted variable bias is the likely explanation for the empirical regularity that the longest lag is largest (demonstrated in Table 2 and also found in the national literature). In particular, for many time series

---

<sup>16</sup> These long-run estimates represent the log-point increase in recipients per capita that would be associated with a one-percentage point increase in the unemployment rate. They are calculated as  $\beta/(1-p)$ , where  $\beta$  is the coefficient on the unemployment rate (or the sum of the unemployment rate and its lags) and  $p$  is the coefficient on the lagged dependent variable. The estimates are difficult to interpret because  $p$  is so large; we present them mainly for completeness.

processes, the omitted variables (i.e., the additional unemployment rate lags) will be most correlated with the last included unemployment rate lag.<sup>17</sup> Adding sufficiently more unemployment rate lags to these models could be problematic, however, because of the relatively short available time series.

In Table 3, we re-estimate the static stock models but include first-order interactions. We estimate these models using the log of the per capita caseload as the dependent variable to facilitate comparisons to the previous results. The interactions are highly jointly significant in all three of the models (F-tests for excluding all interactions produce p-values well under 0.0001). However, our main interest is in the impact of these interactions on the importance of economic conditions. The impact is substantial. For example, by including the interaction between the current and lagged unemployment, the amount of the decline explained by the economy increases from 36.9 percent (see Table 2, column 2) to 71.9 percent (see Table 3, column 1). Similarly, by adding a complete set of first order interactions to the model with two lags, the amount of the decline explained by the economy increases from 35.6 percent (see Table 3, column 3) to 71.5 percent (see Table 3, column 2). The long-run elasticities also increase substantially in each case.

As noted in Section 1, the stock-flow accounting identities suggest that these regressions should be run levels, not logs. Similar regressions in levels (available from the authors on request) show similar patterns. There is a strong sensitivity of results on the number of included lags, the coefficient on the longest lag is largest, and interactions between lags of the unemployment rate enter significantly and increase the implied effect of the economy.

The dynamic stock estimates for the role of the economy are in the 40 percent to 70 percent range (see columns 5 to 8 of Table 2), similarly exhibiting a pattern that the estimates are sensitive to the number of lags that are included. These estimates are suspect for many reasons. On analytic grounds, a stock-flow model suggests that a

---

<sup>17</sup> It is straightforward to show that this claim is true for the case where the truncated explanatory variable follows an AR(1) process. For that case, in fact, it is *only* the last lag that is biased. Pakes and Griliches (1984) provide more general results on truncating lag structures. Interestingly, omitting interactions has the same impact for the case where the lagged independent variable follows an AR(1) process; the intuition for this result is not readily apparent.

lagged dependent variable is only an appropriate reduced-form solution when there is no duration dependence. Such an assumption is clearly rejected by previous studies and by our data. Moreover, such a structural model should include an appropriate interaction and be estimated in levels (see equation 8).

Introducing a lagged dependent variable also causes econometric problems. First, as recognized by previous authors (but not always addressed in estimation), a lagged-dependent variable in combination with the fixed effects identification strategy suggests that the estimates are subject to Nickell bias (Nickell 1981). Importantly, Nickell bias is likely to cause the estimates on the unemployment rate to be too high (assuming that the unemployment rate increases the welfare caseload) and the estimates on the welfare reform variables to be too low (assuming that welfare reform reduced the welfare caseload). In addition, it is well known that lagged dependent variable models are biased if the errors are serially correlated.

Perhaps even more problematic and overlooked by previous authors is that empirical models with a lagged dependent variable face much more difficult "omitted variables bias" problems. In particular, longer-run calculations of the impact of the unemployment rate (such as those presented here) also require an unbiased causal estimate of the coefficient on the lagged dependent variable, implying that any omitted variables must also be uncorrelated with the lagged dependent variable. Such a condition is unlikely because the omitted variables are presumably correlated over time. Consequently, most omitted determinants of the welfare caseload would be proxied for by the lagged dependent variable, resulting in a large coefficient estimate that should not be considered causal; however, the empirical calculations require such an interpretation. Moreover, the estimates in Table 2 imply that it is the coefficient on the lagged dependent variable that drives the larger estimates from the dynamic stock models: the coefficients on the unemployment rate measures are very similar and the coefficient on the lagged dependent variable is near one. Therefore, it is likely that the high estimates produced by these methods are an artifact of the estimation method rather than representing true, large estimates.

#### 4. RESULTS FROM A STOCK-FLOW MODEL

The previous section has shown that the aggregate regressions are consistent with the implications of our stock-flow model. In this section, we directly estimate the stock-flow model and perform simulations to explore its implications for the effect of the economy on the caseload. We begin by developing our methods for estimating the transition probabilities. We then present estimates of the underlying flow models. Finally, we simulate the impact of these flow relationships on the caseload stock for various histories of the unemployment rate.

*Modeling the Flow Relationships.* The key element for the simulation model is the relationship between economic conditions and the flows onto and off of welfare. For analytic convenience, equations 4 and 5 specified a linear relationship between the flows (the continuation rate and the entry rate) and economic conditions. In practice, the flows are bounded between zero and one. Following much of the literature, we estimate the entry rate with a logit specification (e.g., Bane and Ellwood 1986, Blank and Ruggles 1996, and Hoynes 2000). For the continuation rate, we retain the linear specification for computation ease given our very large micro data set (approximately 10 million person-month observations).

We measure economic conditions with various lags of the unemployment rate. We use the unemployment rate, following the aggregate literature, to focus on the impact of using a stock-flow model to examine caseload changes.<sup>18</sup> We include monthly lags of the unemployment rate because of the possibility that lags could have a substantive impact on the decision to enter or remain on welfare: bad economic conditions for multiple periods is likely to be worse than bad economic conditions for only one period. To the extent that significant lagged relationships do not exist in the flow models, then the complicated lag structure that is observed in the stock data can be attributed to the stock-flow process.

---

18 Other authors (Bartik and Eberts, 1999, Ribar, 2000, and Hoynes 2000) have found that richer measures of economic conditions have more explanatory power when examining welfare receipt; thus, our estimates of economic conditions presented here should be considered a lower bound.



Because the MEDS data include information only for those on welfare, we estimate our model for the entry rate (and its dependence on the unemployment rate) using a grouped-data logit model (Maddala, 1983). We calculate the entry rate for county  $j$  in month  $t$ ,  $e_{jt}$ , as the ratio of the number of entrants observed in the MEDS relative to the number of people at risk of going on aid. We then estimate a grouped-data logit model at the county-month level that includes the unemployment rate and fixed effects for time and county,

$$(15) \quad \ln \frac{e_{jt}}{1 - e_{jt}} = \alpha_e + \beta_e Y_{jt} + \gamma_t + \delta_j + \varepsilon_{jt}.$$

Rather than including a full set of dummy variables for each calendar month, we include a piece-wise linear spline (by year and that is not continuous) to capture a general time trends and calendar month dummies (i.e., a dummy for January, February, etc.) to capture seasonal variation.

We estimate the continuation rates  $c^k(Y_t)$  using individual-level data. As noted previously, the continuation rate is simply one minus the conventional hazard rate and thus the estimation problems are equivalent. Let  $C_{ijt}$  be an indicator variable equal to one for whether individual  $i$  in county  $j$  continues on aid in month  $t$ . Consider the model,

$$(16) \quad \Pr[C_{ijt} = 1 | k] = f(\alpha_c + \beta_c Y_{jt} + \gamma_t + \delta_j + g_c(k_{ijt})),$$

where  $g_c(k_{ijt})$  is a flexible specification for the dependence of the continuation probability on  $k_{ijt}$ . Again, we choose  $f$  to be a linear function due to the large data set we use. Finally, we modify the basic specification by adopting a similar time structure as described for the entry rate model, a discontinuous piece-wise annual spline with calendar month dummies.<sup>19</sup>

---

<sup>19</sup> Haider, Klerman, and Roth (2001) extend this model to allow the effect of economic conditions to vary with duration. The interactions are statistically significant, but the basic qualitative findings about the effect of the economy are not affected.

Our primary specification for  $g_c(k_{ijt})$  includes dummy variables for the first six months individuals are on aid and thereafter a quartic in  $k_{ijt}$ . The initial dummy variables will capture the empirical regularity that continuation rates first decline before a systematic increase, captured by the quartic (see Hoynes, 2000).<sup>20</sup>

One complication is that we have data only on current welfare receipt status, and thus, for individuals who receive welfare in the first month of the data, we do not know how long they have received aid. To address this form of left censoring, we assume that the probability of continuation is constant after  $\bar{k}$  periods and then discard the first  $\bar{k}$  periods of the data. Therefore, anyone continuously on aid from the start of the data to period  $\bar{k} + 1$  is in the constant part of the hazard, making the left censoring irrelevant. For everyone else, we know the exact duration. We set  $\bar{k}$  to be 24 months for the results we present here, a specification that is supported by the data.<sup>21</sup> Thus, we drop the monthly data for 1987 and 1988 and use data only for 1989 through 1998 in estimation.

*Estimates with Monthly Data.* We present the results for the entry and continuation models in Tables 4 using monthly data from January 1989 to December 1998, as well as additional results for the static stock model. Each model is estimated with 2, 5, and 11 monthly lags of the unemployment rate.

Focusing on the static stock regressions (columns 1 to 3), we note that the same pattern exists as was present when using annual data (Table 2, columns 1 to 4) with the longest lag always being the largest. It is particularly striking for the model in column 3 that includes 11 monthly lags. In this model, the longest lag is over twice as large as any of the other unemployment rate lags.

Turning to the entry results in columns 4 to 6, the contemporaneous unemployment rate and initial lags are the largest, right-signed and significant for all three of the logistic

---

<sup>20</sup> To arrive at this specification, we estimated models that allowed the monthly pattern to be free by including a complete set of monthly dummy variables. The chosen specification captures the primary features implied by the dummy specification quite well. Further details are available in Haider, Klerman, and Roth (2000).

<sup>21</sup> The results are similar when  $\bar{k}$  is set to 36 months.

regressions.<sup>22</sup> In addition, the impact of the unemployment rate lags tends to decline with the longer lags. Thus, the peculiar results with the stock regressions are not found in entry regressions. Moreover, the results indicate that few *monthly* lags are needed to capture the variation in the data.

The results are just as striking for the continuation rate regressions, also presented in Table 4. Again, current unemployment and the initial lags tend to be large, of the right sign, and statistically significant. In these regressions, the longest lags are an order of magnitude smaller and are not significantly different from zero. Once again, the results indicate that few monthly lags are needed to capture the variation in the data.

These estimates imply that far fewer lags are needed to capture the underlying flow relationships, and thus the complicated lag structure in the stocks is likely due to the stock-flow process itself. Importantly, this finding suggests that far shorter time series will be needed to obtain precise estimates when flow data are used.

*Recovering the Stock Relationships from the Flow Relationships.* The main focus of this paper is the implications of these flow relationships for the responsiveness of the caseload stock to the unemployment rate. To explore that question, we use the flow estimates to calculate the probability that a person transits between the states of being on aid and off aid in each period (i.e., the elements of the transition matrix  $M$ ). These probabilities are then used with the Markov model (equation 12) to simulate how the caseload evolves over time.

We present simulations for the actual history of the unemployment rate path in Figure 2 to demonstrate the fit of the models. We estimate stock-flow models in which we use two, five, and eleven monthly lags for each of the flow relationships, and graph these results along with the actual recipients per capita. While the prediction errors need not be (and are not) mean-zero because these calculations represent a non-linear transformation of the underlying regression estimates (which are mean-zero by construction), the stock-

---

<sup>22</sup> The next sub-section reports simulations to examine the economic significance of these coefficients.

flow model captures the major features of the data. The main difference between the actual recipients per capita and the predicted estimates is that the predicted estimates are too high for the middle of the sample period and decline too quickly towards the end of the sample period.

To obtain estimates of the responsiveness of the caseload to the unemployment rate, we perform simulations for counterfactual (i.e., alternative) histories of the unemployment rate and compare the resulting differences in the caseload. We note from the outset that the richness of our model requires a more precise statement of the counterfactual history than in the conventional regression models. The appropriate counterfactual history depends on the specific question that is being asked.

We present a series of simulations that will provide answers to the question that is usually posed in the literature, "How much of the caseload decline (from its peak) is due to economic conditions?" To answer this question, we compare a simulation based on the actual time and unemployment rate path to a simulation where we assume a different unemployment rate path. The difference between the caseload paths is the impact of the change in the unemployment rate.

We consider two counterfactual histories of the unemployment rate: (1) an unemployment rate path that follows its actual path until the actual caseload peak (March 1995) and then remains constant, and (2) an unemployment rate path that follows its actual path until the actual unemployment rate peak (January 1993) and then remains constant.<sup>23</sup> The first simulation is likely to underestimate the true impact of the economy. In particular, half of the decline in the unemployment rate occurred before the caseload peaked. Specifically, the unemployment rate peaked in January 1993 at 10.3 percent, declined to 7.8 percent when the caseload peaked in March 1995, and further declined to

---

<sup>23</sup> We consider "remains constant" to be the case in which the contemporaneous unemployment rate and its lags remain exactly at the particular date considered. For example, for the counterfactual based around March 1995, the contemporaneous unemployment rate will remain at the March 1995 level and the first lag will remain at the February 1995 level for all future periods. This definition of "remaining constant" is dynamically inconsistent and does not correspond to the description of, "...if the economic conditions in March 1995 persisted forever." However, this counterfactual corresponds most closely to that used by previous authors and presented in earlier subsections (see the appendix). The choice of a dynamically consistent counterfactual has no substantive impact on our results.

5.5 percent through the end of our sample period in December of 1998. Thus, economic conditions would have already put significant downward pressure on the welfare caseload, all else equal, by March 1995. This pressure can be observed in Figure 2 by the significant slowdown of the caseload growth from January 1993 to March 1995. Therefore, we also report simulations using the counterfactual based on holding the unemployment rate fixed at its peak (January 1993) to capture the impact of this earlier employment rate decline. This second simulation is comparable to the earlier stock studies because these studies included much longer lags of the unemployment rate (2 annual lags vs. the 2 to 11 monthly lags we use) and their mis-specified models caused the longest lag to be empirically most important; thus, the earlier studies were also effectively using the decline in the unemployment rate before the caseload peak.

In Figure 3, we present simulations based on the actual unemployment rate for the 2-lag model (because the results for the 5- and 11-lag models are almost identical—see Figure 2) and simulations based on the counterfactual unemployment rate that remains constant after March 1995 for the 2-, 5-, and 11-lag models. As expected, the caseload is higher for each of these counterfactual histories. Comparing the simulated decline to the actual decline (the actual estimates are presented in Table 5), the 2-, 5-, and 11-lag models imply that the caseload would have still declined by 29 percent, 28 percent and 26 percent, respectively. Thus, even without the declining unemployment rate, the caseload would have still declined quite substantially due to the estimated time fixed effects. The fixed effects likely capture unmeasured changes in economic conditions and policy. These models imply that 12 percent, 17 percent and 23 percent of the total caseload decline can be attributed to the declining unemployment rate. This counterfactual, however, ignores the fact that over half of the unemployment decline occurred in the two years before March 1995. The larger impact of the 11-lag model appears to be driven by the substantial declines in the unemployment rate that were occurring between May 1994 and March 1995.

Figure 4 presents a second set of simulations for the counterfactual where the unemployment rate remains constant after January 1993. As is apparent from the figure

(and the actual estimates in Table 5), this counterfactual results in a substantially higher caseload in December 1998, and thus a substantially larger role for the changing unemployment rate. The results imply that the unemployment rate explains 46 percent, 47 percent, and 47 percent for the 2-, 5-, and 11-lag models, respectively (see Table 5). A few interesting points emerge. First, as expected, the caseload is much higher for all models under this counterfactual (compared to those in Figure 3) because the economic conditions are much worse under this scenario, remaining at 10 percent for seven years. Second, although adding lags of up to 11 months does increase the role for the economy, the increase is small. Because of the stability of these estimates, as well as the fact that they capture the entire improvement of the economy, they are the preferred estimates to compare to the previous literature.

## 5. CONCLUSIONS

This paper reconsiders the methods used in the national literature to assess the causes of the recent welfare caseload decline. The literature has produced widely varying conclusions about the causes of the decline and suggested that the varying conclusions arise from differences in the specification of dynamics. We develop an alternative model that is explicitly based on accounting identities implied by the observation that the caseload this period is the result of entry and exit in previous periods. We then present a series of analytical and empirical results that suggest that previous methods are likely to yield biased estimates. Finally, we develop, estimate, and simulate an empirical model based on the underlying flows and provide estimates for the role of the economy in determining California's caseload decline. Our methods suggest that approximately half of the California welfare caseload decline can be attributed to changing economic conditions, as measured by the unemployment rate.

The empirical results provide strong support for the stock-flow approach. California data are sufficiently rich to reproduce the conventional stock results found for the entire United States, and thus they provide a suitable case study to examine our model. Our model provides a plausible explanation for peculiar and sensitive results in the previous

literature, as well as successfully predicts that interactions are likely to be important. Furthermore, while the stock regressions appear to require a large number of lags of the explanatory variables, the flow regressions are nearly a function only of current explanatory variables. Given the short available time-series, this could be a significant advantage for estimation.

The results from this paper have implications beyond the narrow question of the impact of economic conditions on the California welfare caseload. The empirical results for based on California stock regressions are very similar to results presented in the national literature; thus, it is reasonable to expect similar results would be found if national data were available.

This bias in stock regressions equally applies to changes in other potential explanatory variables. For example, consider a one-time policy change that caused the entry rate to decline. Such a change could cause the caseload stock to decline for several years in a stock-flow model. Thus, a conventional differencing identification strategy is likely to underestimate the effect of the policy change, particularly when differencing over short time periods.

Finally, such stock-flow concerns are likely to be important when evaluating many other economic outcomes. For example, Schoeni (2000) demonstrates that the literature examining the changing foodstamp caseload is exhibiting similar patterns (e.g., Wallace and Blank 1999; Currie and Grogger 2000; and Figlio, Gunderson, and Ziliak, and 2000).

## APPENDIX A: Derivations of Propositions and Computational Formulae

**Proposition 2 Proof.** Let  $e_t$  be the entry rate and  $c_t$  be the continuation rate for time  $t$ , and let  $K$  be the maximum number of periods individuals are on aid. If  $K$  is equal to infinity, then the basic accounting identities from the text imply that the caseload can be written as,

$$\begin{aligned}
 n_t &= e_t + c_t e_{t-1} + c_t c_{t-1} e_{t-2} + c_t c_{t-1} c_{t-2} e_{t-3} + \dots \\
 &= e_t + c_t [e_{t-1} + c_{t-1} e_{t-2} + c_{t-1} c_{t-2} e_{t-3} + \dots] \\
 (A1) \quad &= e_t + c_t [n_{t-1}].
 \end{aligned}$$

Assuming that the entry and continuation rates are only functions of contemporaneous economic conditions, we obtain the following expression by substituting in equations 4 and 5 from the text,

$$\begin{aligned}
 n_t &= (\alpha_e + \beta_e Y_t + \varepsilon_{et}) + (\alpha_c + \beta_c Y_t + \varepsilon_{ct}) n_{t-1} \\
 (A2) \quad &= \alpha_e + \beta_e Y_t + \alpha_c n_{t-1} + \beta_c n_{t-1} Y_t + (\varepsilon_{et} + n_{t-1} \varepsilon_{ct}) \\
 &= \theta_1 + \theta_2 Y_t + \theta_3 n_{t-1} + \theta_4 (n_{t-1} Y_t) + v_t.
 \end{aligned}$$

If the lagged caseload is uncorrelated with contemporaneous flow errors (e.g., the errors are serially independent), then only the lagged dependent variable, contemporaneous economic conditions, and a first order interaction are needed to identify the underlying flow parameters.

For most of our results in the text, we assume that  $K$  is finite. In this case, the general result goes through only up to an approximation. In particular, A1 becomes

$$\begin{aligned}
 n_t &= e_t + c_t e_{t-1} + c_t c_{t-1} e_{t-2} + \dots + c_t c_{t-1} \dots c_{t-K+1} e_{t-K} \\
 (A3) \quad &= e_t + c_t [e_{t-1} + c_{t-1} e_{t-2} + \dots + c_{t-1} \dots c_{t-K+1} e_{t-K}] \\
 &= e_t + c_t [n_{t-1} - c_{t-1} \dots c_{t-K} e_{t-K-1}] \\
 &\doteq e_t + c_t n_{t-1}.
 \end{aligned}$$



The quality of the approximation in the last step will depend on the share of the caseload that is on aid for period  $K$ , which declines geometrically in  $K$ .

**Proposition 3 Proof.** Let  $e_t$  be the entry rate for period  $t$  and  $c_t^k$  be the continuation probability for individuals on aid for  $k$  periods in period  $t$ . Using the basic accounting identities from the text, the caseload can be written as,

$$\begin{aligned}
 n_t &= e_t + c_t^1 e_{t-1} + c_t^2 c_{t-1}^1 e_{t-2} + \dots + c_t^K c_{t-1}^{K-1} \dots c_{t-K+1}^1 e_{t-K} \\
 (A4) \quad &= e_t + \left[ c_t^1 \frac{e_{t-1}}{n_{t-1}} + c_t^2 \frac{c_{t-1}^1 e_{t-2}}{n_{t-1}} + \dots + c_t^K \frac{c_{t-1}^{K-1} \dots c_{t-K+1}^1 e_{t-K}}{n_{t-1}} \right] n_{t-1} \\
 &= e_t + [c_t^1 \lambda_{t-1}^1 + c_t^2 \lambda_{t-1}^2 + \dots + c_t^K \lambda_{t-1}^K] n_{t-1}.
 \end{aligned}$$

where  $\lambda_t^k$  is the share of the caseload in period  $t$  that began  $k$  periods earlier. The weights are a function only of past economic conditions.

If the entry and continuation rates are functions only of contemporaneous economic conditions, then A4 can be written as follows

$$\begin{aligned}
 n_t &= (\alpha_e + \beta_e Y_t + \varepsilon_{et}) + n_{t-1} \sum_{k=1}^K (\alpha_c^k + \beta_c^k Y_t + \varepsilon_{ct}^k) \lambda_{t-1}^k \\
 (A5) \quad &= \alpha_e + \beta_e Y_t + \left( \sum_{k=1}^K \alpha_c^k \lambda_{t-1}^k \right) n_{t-1} + \left( \sum_{k=1}^K \beta_c^k \lambda_{t-1}^k \right) Y_t n_{t-1} + \left( \varepsilon_{et} + n_{t-1} \sum_{k=1}^K \lambda_{t-1}^k \varepsilon_{ct}^k \right)
 \end{aligned}$$

where  $k$  indexes the duration-specific parameters for the continuation probabilities (i.e.,  $\alpha_c^k$  and  $\beta_c^k$ ). This expression implies that the coefficients on  $n_{t-1}$  and  $Y_t n_{t-1}$  are a weighted average of the duration-specific continuation parameters.

Now, consider the weights derived above,

$$\begin{aligned}
 \lambda_t^k &= \frac{e_{t-k}}{n_{t-1}} \prod_{j=1}^{k-1} c_{t-(K+j)}^j \\
 (A6) \quad &= \frac{(\alpha_e + \beta_e Y_{t-k} + \varepsilon_{e,t-k})}{n_{t-1}} \prod_{j=1}^{k-1} (\alpha_c^j + \beta_c^j Y_{t-(K-j)} + \varepsilon_{c,t-(K-j)}^j), \quad \forall k = 2, 3, \dots, K.
 \end{aligned}$$

Note that because  $n_{t-1}$  depends on  $K$  periods of economic conditions, the weights depend on  $K$  periods of economic conditions. In addition, the weights change in a predictable manner with previous economic conditions. To see this, consider the ratio of the weights for durations  $k+1$  and  $k$ ,

$$(A7) \quad \frac{\lambda_t^{k+1}}{\lambda_t^k} = \frac{e_{t-(k+1)} c_t^k}{e_{t-k}} = \frac{e_{t-(k+1)} (\alpha_c^k + \beta_c^k Y_{t-k} + \varepsilon_{ct}^k)}{\alpha_c + \beta_c Y_{t-k} + \varepsilon_{c,t-k}}$$

If the economy were to improve in period  $t-k$ , this ratio would change depending on the relative responsiveness of the continuation probability for duration  $k$ , weighted by the number of entrants in the previous period, and the entry rate in period  $t-k$ . Thus, the relative weights depend on the history of economic conditions.

**Calculating the Impact of Economic Conditions.** Although such calculations are straightforward for the static case with no lags, they are not as clear when independent and dependent lags are included. Therefore, we explicitly develop general notation for the simple case, so that our extensions are explicit.

Define  $\tilde{n}(Y = a, t = b)$  to be the predicted welfare recipients per capita based on the coefficients of equation 11 with the unemployment rate for year  $a$  and the time effects associated with year  $b$ ,

$$(A8) \quad \begin{aligned} \tilde{n}(Y = a, t = b) &= E[\exp[\hat{\alpha} + Y_a \hat{\beta} + \hat{\gamma}_b + e_b]] \\ &= \exp[\hat{\alpha} + Y_a \hat{\beta} + \hat{\gamma}_b] E[\exp[e_b]] \end{aligned}$$

Then, we define the impact of economic conditions to be

$$(A9) \quad \begin{aligned} \Delta_{econ}^S &= \frac{[\tilde{n}(Y = 1998, t = 1994) - \tilde{n}(Y = 1994, t = 1994)] / \tilde{n}(Y = 1994, t = 1994)}{(n_{1998} - n_{1994}) / n_{1994}} \\ &= \frac{\exp[\hat{\beta}(Y_{1998} - Y_{1994})] - 1}{(n_{1998} - n_{1994}) / n_{1994}} \end{aligned}$$

This expression is readily extended to include additional measures of economic conditions such as the lagged unemployment rate by considering  $Y$  and its coefficient to be vector valued

To develop a similar expression for the dynamic stock model (i.e., including a lagged dependent variable), we begin by noting that the expression inside the exponential function of equation A8 is simply the change in the log-per capita model due to economic conditions. We proceed analogously for the dynamic model. Consider the change in the unemployment rate from 1994 to 1995. For 1995, this change is associated with a  $\hat{\beta}(Y_{1995} - Y_{1994})$  increase, just as it would for the static stock model case. However, due to the presence of a lagged dependent variable, this unemployment rate change will have additional impact between 1995 and 1996, of  $\hat{\beta}(Y_{1995} - Y_{1994})\hat{\rho}$ . Using a similar argument for all of the other yearly changes, the expression for the dynamic model becomes

$$(A10) \quad \Delta_{econ}^D = \frac{\exp[F] - 1}{(n_{1998} - n_{1994}) / n_{1994}}$$

$$\text{where} \quad F = \hat{\beta}(1 + \hat{\rho} + \hat{\rho}^2 + \hat{\rho}^3)(Y_{1995} - Y_{1994}) + \hat{\beta}(1 + \hat{\rho} + \hat{\rho}^2)(Y_{1996} - Y_{1995}) + \dots + \hat{\beta}(Y_{1998} - Y_{1997})$$

Again, this expression can be extended for additional measures of economic conditions such as lagged economic conditions by simply repeating this expression for the other included variables.

## APPENDIX B: The MEDS Data

The methodology we apply in the body of the paper exploits the availability of detailed longitudinal information on individuals in the welfare caseload to identify flows onto and off of welfare. We apply the methodology to administrative data on the welfare caseload for the state of California that is extracted from the Medi-Cal Eligibility Determination System (MEDS). The MEDS is a monthly roster of all individuals eligible

for Medi-Cal, California's Medicaid program, that is used for administrative purposes. Because welfare recipients are categorically eligible for Medi-Cal and such eligibility is noted, MEDS provides a monthly roster of the welfare population in California. See Haider et al (2000) for more details concerning the MEDS.

We construct an analysis file from the MEDS by drawing a stratified random sample of all individuals on aid, such that we obtain approximately a 3 percent sample of the population. We stratify by county and draw a random sample so that we have approximately an equal numbers of observations for each of California's 58 counties. This scheme results in an analysis file that contains 282,381 people who have ever received cash assistance, comprising 487,464 spells and 10,966,420 person-months of cash aid receipt during the years 1989 to 1998 (our eventual sample period). To avoid some small sample problems, we aggregate California's five smallest counties into a single "county" for analysis purposes. The five smallest counties are Alpine, Colusa, Modoc, Mono, and Sierra; combined, their welfare population represents well under one percent of the state's welfare population. We perform all of our analyses on these 53 counties and 1 county group.

Previous research indicates that there is considerable "churning" on and off welfare in the MEDS data. This churning is likely due to administrative record keeping rather than "real" entrances and exits (see Hoynes, 2000). To mitigate such concerns, we recode one-month spells on and off of aid as not having occurred, following Hoynes (2000).

In addition to the MEDS, we also rely on various data sets that are publicly available. We use the CA237, the official monthly caseload reports from the counties to the California Department of Social Services; these data are described in Haider, et al, (2000). We use Intercensal Population Estimates for each county, generated by the U.S. Bureau of the Census. For all of the estimates presented in this paper, we consider the population at risk of being on aid to be everyone under the age of 50. The population estimates are only available by year, so we perform a simple linear interpolation to obtain monthly data. Finally, we use local area unemployment estimates produced by the U.S. Department of Labor to proxy for economic conditions.

## REFERENCES

- Bane, Mary Jo and David Ellwood (1986). "Slipping Into and Out of Poverty: The Dynamics of Spells." *Journal of Human Resources*; 21(1), Winter 1996, 1-23.
- Bane, Mary Jo and David Ellwood (1994). *Welfare Realities: From Rhetoric to Reform: Preface*. Cambridge, MA: Harvard University Press.
- Bartik, Timothy and Randall Eberts (1999). "Examining the Effect of Industry Trends and Structure on Welfare Caseloads." In *Economic Conditions and Welfare Reform* (ed. Sheldon Danziger), Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.
- Blank, Rebecca (1989). "Analyzing the Length of Welfare Spells." *Journal of Public Economics*, 39, 245-273.
- Blank, Rebecca and Patricia Ruggles (1996). "When Do Women Use AFDC and Food Stamps? The Dynamics of Eligibility vs. Participation." *Journal of Human Resources*, 31:1, 57-89.
- Congressional Budget Office (1993). "Forecasting AFDC Caseloads, with an Emphasis on Economic Factors." Staff memorandum. July.
- Council of Economic Advisers (1997). "Technical Report: Explaining the Decline in Welfare Receipt, 1993-96." A Report by the Council of Economic Advisers, May.
- Council of Economic Advisers (1999). "Technical Report: The Effects of Welfare Policy and the Economic Expansion on Welfare Caseloads: An Update." A Report by the Council of Economic Advisors, August.
- Currie, Janet and Jeff Grogger (2000). "Explaining Recent Declines in the Food Stamp Program Participation." September manuscript.
- Figlio, David and James Ziliak (1999). "Welfare Reform, the Business Cycle, and the Decline in AFDC Caseloads." In *Economic Conditions and Welfare Reform* (ed. Sheldon Danziger), Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.
- Figlio, D. N., Gunderson, C., & Ziliak, J. P. (2000). "The effects of the macroeconomy and welfare reform on food stamp caseloads." January 2000.
- Gittleman, Maury (1998). "Declining Caseloads: What Do the Dynamics of Welfare Participation Reveal?" December 1998 manuscript.
- Haider, Steven, Jacob Alex Klerman, Jan Hanley, Laurie McDonald, Elizabeth Roth, Liisa Hiatt, and Marika Suttorp (1999). "Welfare Reform in California: Design of

the Impact Analysis, Preliminary Investigations of Caseload Data." RAND DRR(L)-2077/1-CDSS, May.

Haider, Steven, Jacob Alex Klerman, and Elizabeth Roth (2000). "Changes in the Welfare Caseload: How Do Economic Conditions Matter?" October 2000 manuscript.

Heckman, James and James Walker (1990). "The Relationship between Wages and income and the Timing and Spacing of Births: Evidence from Swedish Longitudinal Data." *Econometrica*, 58(6) (Nov.), 1411-1441.

Hoynes, Hilary Williamson (2000). "Local Labor Markets and Welfare Spells: Do Demand Conditions Matter?" *Review of Economics and Statistics*, 82(3), 351-368.

Hoynes, Hilary Williamson and Thomas MaCurdy (1994). "Has the Decline in Benefits Shortened Welfare Spells?" *American Economic Review*, 84(2) 43-48.

Hutchens, Robert (1981). "Entry and Exit Transitions in a Government Transfer Program: The Case of Aid to Families with Dependent Children." *Journal of Human Resources*, 16(2).

Levine, Phillip and Diane Whitemore (1998). "The Impact of Welfare Reform on the AFDC Caseload" *National Tax Association Proceedings*. Washington, DC: Nation Tax Association, 24-33.

Lippman, S. and J. McCall (1976). "The Economics of Job Search: A Survey" *Economic Inquiry*. 14:155-189, 347-368.

Maddala, G. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

MaCurdy, Thomas, David Mancuso, and Margaret O'Brien-Strain (2000). "The Rise and Fall of California's Welfare Caseload: Types and Regions, 1980-99." *Public Policy Institute of California Report*, June.

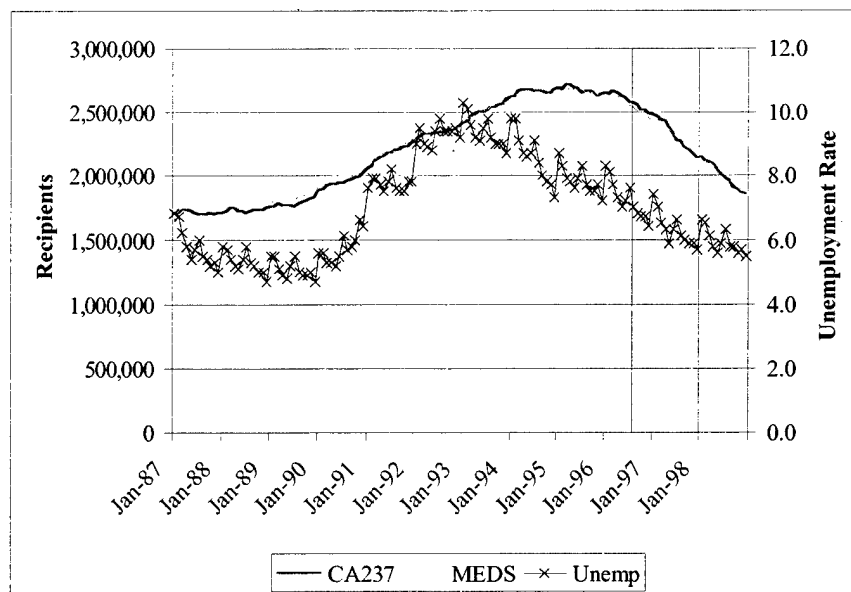
Moffitt, Robert (1983). "An Economic Model of Welfare Stigma." *American Economic Review* 73 (5):1023-1035.

Moffitt, Robert (1992). "Incentive Effects of the U.S. Welfare System: A Review." *Journal of Economic Literature*, 30, 1-61.

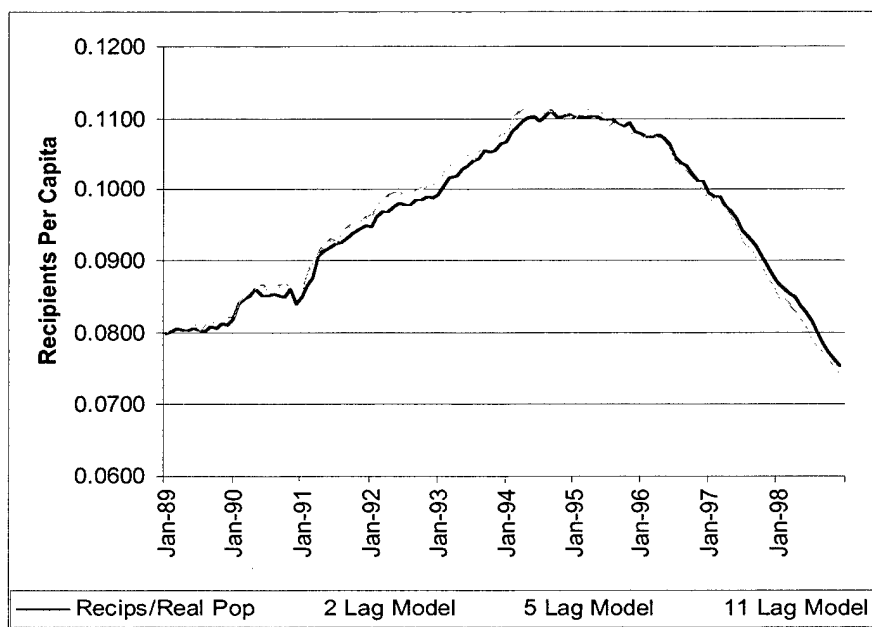
Moffitt, Robert (1993). "Identification and Estimation of Dynamic Models with a Time Series of Repeated Cross-Sections." *Journal of Econometrics*, 59, 99 – 123.

Moffitt, Robert (1999). "The Effect of Pre-PRWORA Waivers on AFDC Caseloads and Female Earnings, Income, and Labor Force Behavior." May manuscript.

- Moffitt, Robert and Michael Rendall (1995). "Cohort Trends in the Lifetime Distribution of Female Family Headship in the United States, 1968-1985." *Demography*, 32(3) (Aug.), 407-424.
- Nickell, Stephen (1981). "Biases in Dynamic Models with Fixed Effects." *Econometrica*, 49(6) (Nov.), 1417-1426.
- Pakes, Ariel and Zvi Griliches (1984). "Estimating Distributed Lags in Short Panels with an Application to the Specification of Depreciation Patterns and Capital Stock Constructs." *Review of Economic Studies*, 51(2) 243-262.
- Ribar, David (2000). "Transitions from Welfare and the Employment Prospects of Low-skill Workers." March manuscript.
- Schoeni, Robert (2000). "Comment on 'Explaining Recent Declines in Food Stamp Program Participation, Currie and Grogger 2000'." September manuscript.
- Ziliak, James, David Figlio, Elizabeth Davis, Laura Connolly (2000). "Accounting for the Decline in AFDC Caseloads: Welfare Reform or Economic Growth?" *Journal of Human Resources*, 35(3), 570 - 586.

**Figure 1: Welfare Recipients and Unemployment Rate in California**

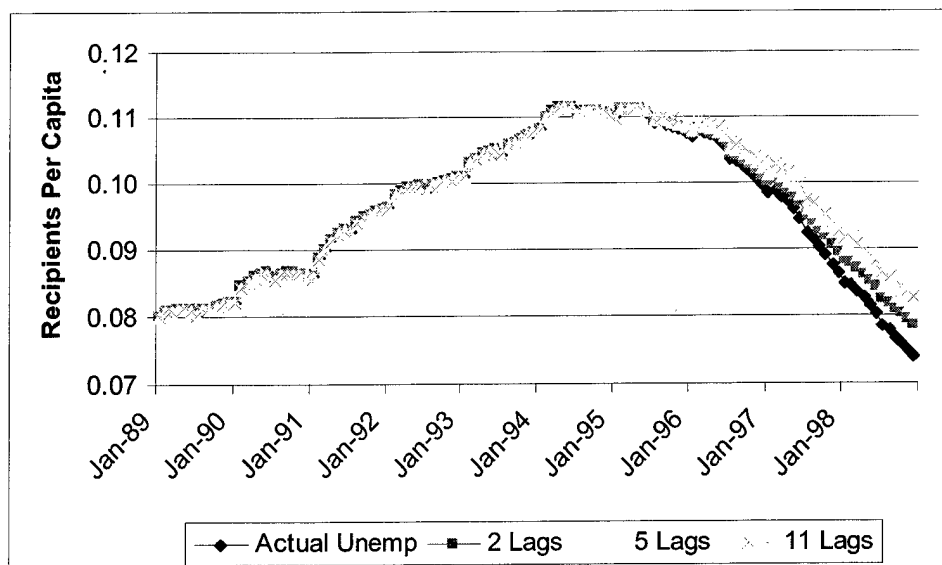
Note: Authors' tabulations from the MEDS, CA237, and the California unemployment rate. The CA237 is the official California welfare caseload. The MEDS represents an estimate of the caseload based on the 3% sample we analyze in this paper. The first vertical line represents in the passage of the federal welfare reform legislation (PRWORA, August 1996) and the second vertical line represents the implementation of the California welfare reform (CalWORKs, January 1998).

**Figure 2: The Actual and Simulated Per Capita Welfare Caseload**

Note: Authors' tabulations from the MEDS. The lower line (with x's) represents the real recipients per capita in California. The other three lines, which lie on top of each other, represent the simulated recipients per capita using the stock-flow model. We present results for the model that includes 2, 5, and 11 lags of the unemployment rate.

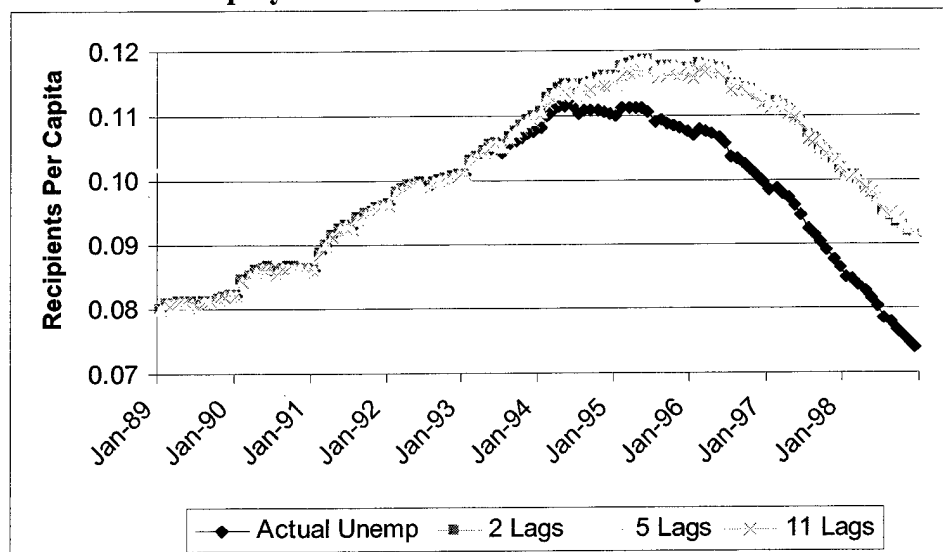


**Figure 3: The Simulated Welfare Caseload in California--Holding the Unemployment Rate Constant after March 1995**



Note: Authors' tabulations from the MEDS. All estimates are based on the simulation model, including the estimated time effects. The four lines correspond to simulation the number of recipients per capita based on the actual unemployment rate (using 2 lags in both flow equations) and based on holding the unemployment rate (and its lags) constant after March 1995 using 2 lags, 5 lags, and 11 lags.

**Figure 4: The Simulated Welfare Caseload in California--Holding the Unemployment Rate Constant in January 1993**



Note: Authors' tabulations from the MEDS. The four lines correspond to simulation the number of recipients per capita based on the actual unemployment rate (using 2 lags in both flow equations) and based on holding the unemployment rate (and its lags) constant after January 1993 using 2 lags, 5 lags, and 11 lags.

**Table 1: Short Term Spell Durations, 1989-1997**

<i>Spell Start Period</i>	<i>Spells</i>	<i>Average Monthly Entry Rate</i>	<i>Average Monthly Continuation Rate for Spells that Lasted:</i>			
			<i>2-5 mos.</i>	<i>6-11 mos.</i>	<i>12-17 mos.</i>	<i>18+ mos.</i>
1/89-12/90	70,721	0.0032	0.938	0.943	0.964	0.979
1/91-12/92	79,620	0.0037	0.942	0.950	0.968	0.982
1/93-12/94	80,863	0.0037	0.946	0.951	0.968	0.982
1/95-12/96	72,234	0.0031	0.942	0.949	0.962	0.977
1/97-12/97**	29,862	0.0024	0.933	0.942	X	X

Note: Authors' tabulations from the MEDS data. The entry rate is calculated as the total number of entrants as a proportion of the population under age 50. An "X" indicates that the probability could not be calculated because of right-censoring. The entry rate represents that average monthly entry rate.

**Table 2: Welfare Stock Regressions Using Annual Data**

<i>Regressors</i>	<i>Log(Recipients Per Capita)</i>				<i>Log(Recipients Per Capita)</i>			
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>
Unemp. Rate	0.022 (0.007)	-0.001 (0.009)	-0.008 (0.009)	-0.003 (0.009)	0.028 (0.003)	0.019 (0.004)	0.018 (0.004)	0.020 (0.004)
Unemp. Rate-1 <sup>st</sup> Lag		0.035 (0.009)	0.013 (0.010)	0.006 (0.010)		0.013 (0.004)	0.009 (0.005)	0.006 (0.005)
Unemp. Rate-2 <sup>nd</sup> Lag			0.042 (0.008)	0.026 (0.009)			0.007 (0.004)	0.001 (0.005)
Unemp. Rate-3 <sup>rd</sup> Lag				0.030 (0.008)				0.014 (0.004)
Dep. Variable-1 <sup>st</sup> Lag					0.976 (0.023)	0.967 (0.023)	0.958 (0.023)	0.951 (0.023)
R-squared	0.9676	0.9686	0.9700	0.9707	0.9927	0.9928	0.9929	0.9930
<i>Impact of Unemp. Rate</i>								
% Attributable of 94 - 98 Decline	20.5%	36.9%	35.6%	20.4%	61.8%	72.9%	69.8%	41.1%
Long-run elasticity	0.022	0.035	0.047	0.059	1.17*	0.99*	0.84*	0.82*

Note: Authors' tabulations from the MEDS data. All models are estimated on data aggregated to the county/year level, include county and time fixed effects, and use data for 1988-1998. Standard errors are in parentheses. The long-run elasticity is the log-point impact of a permanent change in the unemployment rate; see the text for further details. \*These long-run elasticities are difficult to interpret and are only reported for completeness.

**Table 3: Welfare Stock Regressions Using Annual Data**

<i>Regressors</i>	<i>Log(Recipients Per Capita)</i>		
	<i>1</i>	<i>2</i>	<i>3</i>
Unemp. Rate	0.050 (0.009)	0.019 (0.015)	0.026 (0.016)
Unemp. Rate-1 <sup>st</sup> lag	0.093 (0.009)	0.074 (0.018)	0.051 (0.019)
Unemp. Rate-2 <sup>nd</sup> lag		0.075 (0.013)	0.079 (0.018)
Unemp. Rate-3 <sup>rd</sup> lag			0.029 (0.014)
Int(unemp, 1 <sup>st</sup> lag)	-0.004 (0.0004)	-0.002 (0.0008)	-0.008 (0.002)
Int(unemp, 2 <sup>nd</sup> lag)		-0.0002 (0.001)	0.002 (0.004)
Int(unemp, 3 <sup>rd</sup> lag)			0.004 (0.004)
Int(1 <sup>st</sup> lag, 2 <sup>nd</sup> lag)		-0.003 (0.001)	0.002 (0.004)
Int(1 <sup>st</sup> lag, 3 <sup>rd</sup> lag)			0.003 (0.005)
Int(2 <sup>nd</sup> lag, 3 <sup>rd</sup> lag)			-0.008 (0.002)
R-squared	0.9752	0.9772	0.9781
<i>Impact of Unemp. Rate</i>			
% Attributable of 94 – 98 Decline	71.9%	71.5%	66.7%
Long-run elasticity	0.079	0.095	0.104

Note: Authors' tabulations from the MEDS data. All models are estimated on data aggregated to the county/year level, include county and time fixed effects, and use data for 1988-1998. Standard errors are in parentheses. Int(A, B) represents an interaction between A and B. To evaluate the long-run elasticity, we assume that the unemployment rate changes from 7 to 8 percent.

**Table 4: Stock, Entry and Continuation Regressions Using Monthly Data**

<i>Regressors</i>	<i>Static Stock Regressions (OLS)</i>			<i>Probability of Entry (Grouped Logit)</i>			<i>Probability of Continuation (OLS: coeffs &amp; s.d. * 100)</i>		
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>
Unemp.	0.003	0.007	-0.006	0.030	0.036	0.022	0.121	0.115	0.097
Rate	(0.002)	(0.002)	(0.002)	(0.006)	(0.006)	(0.007)	(0.009)	(0.009)	(0.011)
UR-1 <sup>st</sup>	-0.001	0.001	0.006	0.019	0.024	0.031	-0.017	-0.034	-0.037
Lag	(0.003)	(0.003)	(0.003)	(0.008)	(0.009)	(0.009)	(0.013)	(0.014)	(0.014)
UR-2 <sup>nd</sup>	0.008	0.001	0.002	-0.020	-0.040	-0.040	-0.003	0.065	0.081
Lag	(0.002)	(0.003)	(0.003)	(0.006)	(0.009)	(0.009)	(0.009)	(0.014)	(0.014)
UR-3 <sup>rd</sup>		0.003	0.003		0.010	0.013		-0.081	-0.078
Lag		(0.003)	(0.003)		(0.009)	(0.009)		(0.014)	(0.014)
UR-4 <sup>th</sup>		-0.002	0.001		0.006	0.017		0.049	0.038
Lag		(0.003)	(0.003)		(0.009)	(0.009)		(0.014)	(0.014)
UR-5 <sup>th</sup>		0.009	0.005		0.006	-0.013		-0.026	-0.002
Lag		(0.002)	(0.003)		(0.006)	(0.009)		(0.010)	(0.014)
UR-6 <sup>th</sup>			0.001			0.014			-0.030
Lag			(0.003)			(0.009)			(0.014)
UR-7 <sup>th</sup>			0.002			0.001			-0.027
Lag			(0.003)			(0.009)			(0.014)
UR-8 <sup>th</sup>			0.001			-0.008			0.064
Lag			(0.003)			(0.009)			(0.014)
UR-9 <sup>th</sup>			0.001			0.009			-0.016
Lag			(0.003)			(0.008)			(0.014)
UR-10 <sup>th</sup>			-0.001			0.012			-0.020
Lag			(0.003)			(0.008)			(0.014)
UR-11 <sup>th</sup>			0.013			0.003			0.023
Lag			(0.002)			(0.007)			(0.011)
R-squared	0.9669	0.9671	0.9677	0.8216	0.8222	0.8232	0.0096	0.0096	0.0096

Note: Authors' tabulations from the MEDS. We use monthly data for the period January 1989 to December 1998 for each set of models, with the lags referring to monthly lags. All models contain county fixed effects and a flexible spline in time. See the text for further details. Standard errors are in parentheses.

**Table 5: Simulation Results for the Per Capita Caseload**

	<i>Percentage of Mar 95 to Dec 98 Decline Due to Economic Conditions</i>		
	2 Lag Model	5 Lag Model	11 Lag Model
<i>Simulations with Actual Unemp. Rate</i>			
Simulated March 1995 level	0.107	0.107	0.107
Simulated December 1998 level	0.071	0.071	0.071
Simulated percent decline	-33.5%	-33.4%	-33.4%
<i>Simulations to Compare to Previous Methods</i>			
(1) Unemp. rate remains constant after caseload peak (3/95)			
Simulated December 1998 level	0.075	0.077	0.079
Simulated percent decline	-29.4%	-27.7%	-25.6%
Decline attributable to economic conditions	12.2%	17.1%	23.3%
(2) Unemp. rate remains constant after unemployment rate peak (1/93)			
Simulated December 1998 level	0.087	0.088	0.088
Simulated percent decline	-18.1%	-17.7%	-17.5%
Decline Attributable to economic conditions	46.1%	47.0%	47.4%

Note: Authors' tabulations from the MEDS. Simulations are based on the stock-flow model developed in equations 14 to 18. All calculations are based on monthly data for the period January 1989 to December 1998. See the text further details.